



Argonne
NATIONAL
LABORATORY

... for a brighter future



U.S. Department
of Energy

UChicago ►
Argonne_{LLC}



U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC

Electronic Structure Calculations Using Nonlinear Basis Expansions

*Ron Shepard and Gergely Gidofalvi
Chemical Sciences and Engineering Division
Argonne National Laboratory*

Special Session in Honor of Russell M. Pitzer

*63rd OSU International Symposium on Molecular Spectroscopy,
June 16-20, 2008*

Background: Graphical Unitary Group Approach

$$N = 2a + b$$

$$S = b / 2$$

$$n = a + b + c$$

For Full-CI Expansions

$$N_{csf} = \frac{b+1}{n+1} \binom{n+1}{a} \binom{n+1}{c}$$

J. Paldus, *J. Chem. Phys.* **61**, 5321 (1974)

$$N_{row} = (a+1)(b+1)(c+1) + \frac{1}{6}\eta(\eta+1)(2\eta+3|a-c|+1)$$

I. Shavitt, *IJQC* **S11**, 131 (1977)

$$N_{arc} = (2a+1)b(2c+1) + \frac{1}{3}\eta(\eta+1)(4\eta+5) + |a-c|(\eta+1)(2\eta+1)$$

$$N_{\phi} = b(3ac + a + c) - (ac + a + c) + \frac{1}{2}(\eta+1)[\eta(2\eta+3) + |a-c|(3\eta+2)]$$

R. Shepard and M. Minkoff, *IJQC* **106**, 3190 (2006)

$$N_{pair}^{unique} = 10\left(abc + \frac{1}{6}\eta(\eta+1)(2\eta+3|a-c|+1)\right) - 3ac + (a+c-2)(3b+\delta_{0b})$$

$$+ 5b + 3 + (1-\delta_{0b})(\delta_{0c}(ab+b-1) + \delta_{0a}(cb+b-1))$$

$$\eta = \text{Min}(a, c)$$

R. Shepard, M. Minkoff, and S. R. Brozell, *IJQC* **107**, 3203 (2007)

GUGA...

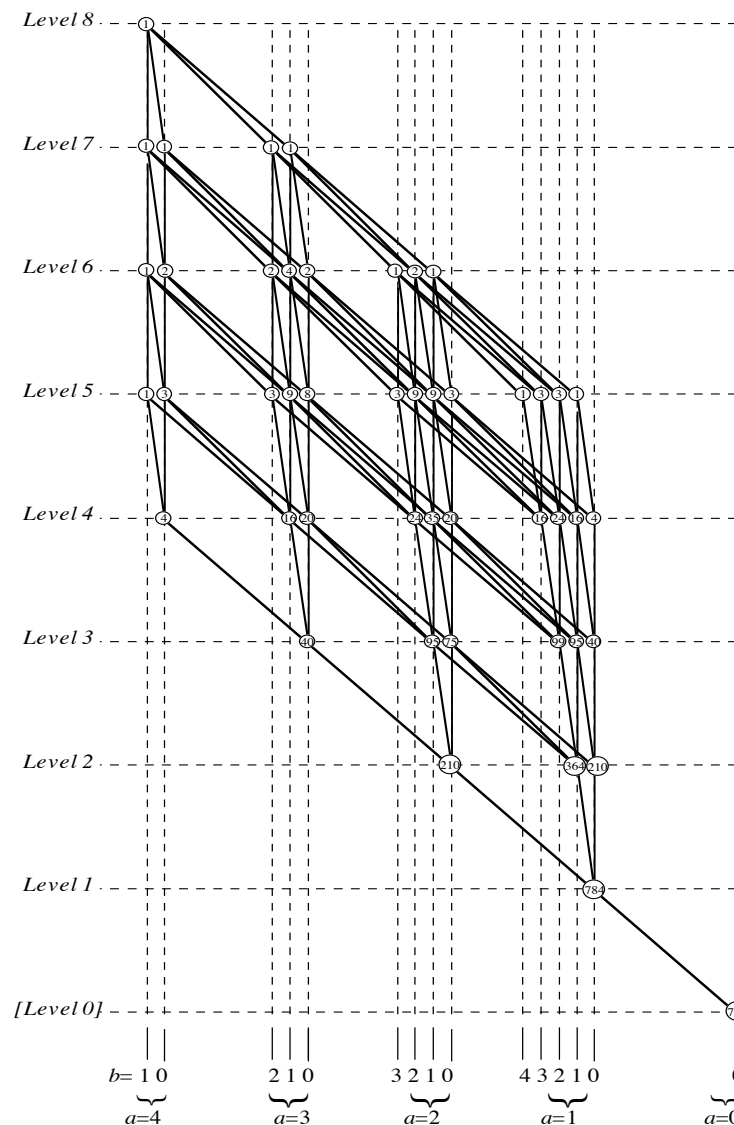
Example:

Shavitt graph for the (1²⁷⁷) CASSCF Expansion Space for CH₃.

This wave function expansion is sufficiently flexible to allow dissociation to any of the ground or valence state fragments: $\text{CH}_2(^3\text{B}_1)$, $\text{H}(^2\text{S})$, $\text{H}_2(^1\Sigma_g)$, $\text{CH}(^2\Pi)$, and $\text{C}(^3\text{P})$.

This expansion space has 784 CSFs; the DRT has 47 Distinct Rows and 118 arcs. (a,b,c)=(4,1,3).

An individual arc or node might be shared by many walks.



GUGA...

For Full-CI singlet expansions with $N=n$ for large n

$$N_{csf} = \frac{1}{n+1} \binom{n+1}{\frac{1}{2}n} \approx \left(\frac{8}{\pi}\right) \frac{4^n}{n^2}$$

$$N_{row} = (n+2)(n+3)(n+4) / 24 \approx \frac{1}{24} n^3$$

$$N_{arc} = n(n+2)(2n+5) / 12 \approx \frac{1}{6} n^3 \approx 4N_{row}$$

$$N_{\phi} = n(n^2 + 3n - 2) / 8 \approx \frac{1}{8} n^3 \approx 3N_{row}$$

$$n \approx \text{Log}_4(N_{csf})$$

$$N_{row} \approx \frac{1}{24} \text{Log}_4^3(N_{csf})$$

GUGA...

Full-CI (Singlet, $n=N$)

$n=N$	N_{csf}	N_{row}
2	3	5
4	20	14
6	175	30
8	1,764	55
10	19,404	91
12	226,512	140
14	2,760,615	204
16	34,763,300	285
18	449,141,836	385
20	5,924,217,936	506
22	79,483,257,308	650
24	1,081,724,803,600	819
26	14,901,311,070,000	1015
28	207,426,250,094,400	1240
30	2,913,690,606,794,775	1496
32	41,255,439,318,353,700	1785
34	588,272,005,095,043,500	2109
36	8,441,132,926,294,530,000	2470
38	121,805,548,126,430,067,900	2870
40	1,766,594,752,418,700,032,400	3311
42	25,739,723,541,439,406,257,200	3795
44	376,607,675,256,599,252,232,000	4324
46	5,531,425,230,331,301,517,157,500	4900

Graphically Contracted Wave Function Approach

- In the Graphically Contracted Wave Function (GCF) approach, each arc in the Shavitt Graph is associated with an **arc factor** $\alpha_{d,j}$
- The linear expansion coefficient associated with the walk m has the value x_m determined as the product of the arc factors

$$m = 1 + \sum_{p=0}^{n-1} y_{d_p j_p}$$

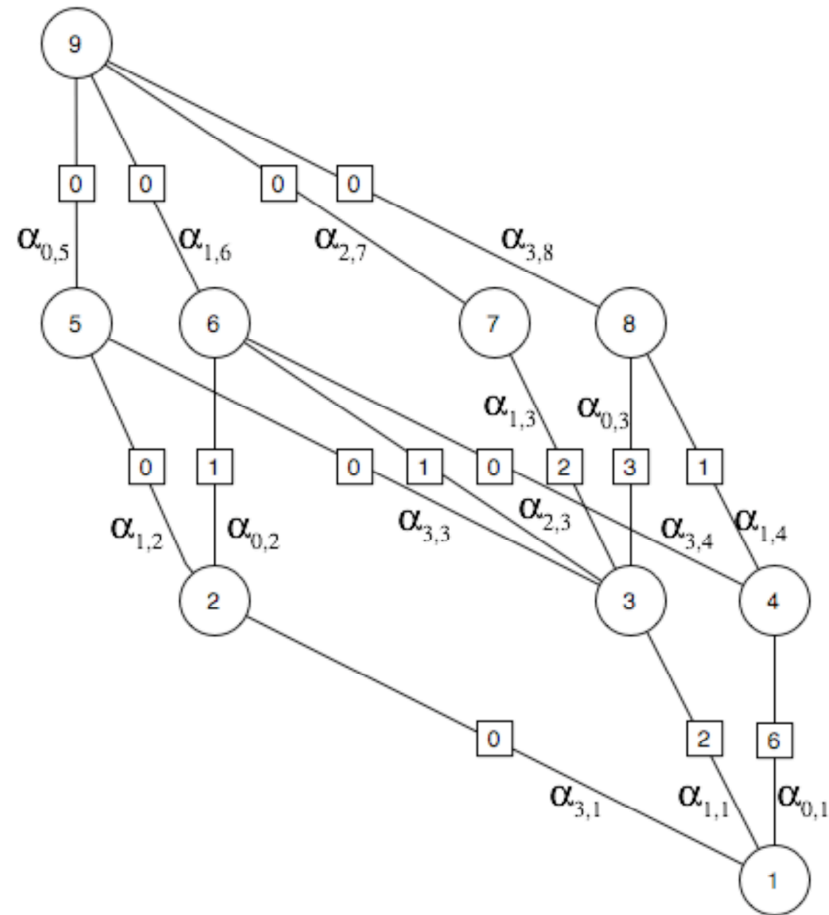
$$x_m = \prod_{p=0}^{n-1} \alpha_{d_p j_p}$$

$$|P\rangle = \sum_m x_m |\tilde{m}\rangle$$

GCF...

Example: 3^3 doublet Full-CI

$$\begin{aligned}
 |P\rangle &= \sum_{m=1}^8 x_m |\tilde{m}\rangle \\
 &= (\alpha_{3,1} \alpha_{1,2} \alpha_{0,5}) |310\rangle \\
 &+ (\alpha_{3,1} \alpha_{0,2} \alpha_{1,6}) |301\rangle \\
 &+ (\alpha_{1,1} \alpha_{3,3} \alpha_{0,5}) |130\rangle \\
 &+ (\alpha_{1,1} \alpha_{2,3} \alpha_{1,6}) |121\rangle \\
 &+ (\alpha_{1,1} \alpha_{1,3} \alpha_{2,7}) |112\rangle \\
 &+ (\alpha_{1,1} \alpha_{0,3} \alpha_{3,8}) |103\rangle \\
 &+ (\alpha_{0,1} \alpha_{3,4} \alpha_{1,6}) |031\rangle \\
 &+ (\alpha_{0,1} \alpha_{1,4} \alpha_{3,8}) |013\rangle
 \end{aligned}$$



$b =$	1	0	2	1	0
$a =$		1			0

GCF-CI...

- A single contracted function $|P\rangle$, corresponding to a set $\{\alpha^P\}$, is similar to other nonlinear expansion methods (PP-GVB, CCSD, DMRG, geminal product wfns) and to internal/external contracted-CI wfns.
- In our GCF-CI approach, the wave function is a **linear combination** of several GCF expansion terms.

$$|\psi\rangle = \sum_P^{N_\alpha} c_P |P\rangle$$

- The expansion coefficients c_P are computed variationally from the generalized real symmetric eigenvalue equation

$$\mathbf{H} \mathbf{c} = \mathbf{S} \mathbf{c} E$$

with $H_{PQ} = \langle P | H | Q \rangle$, $S_{PQ} = \langle P | Q \rangle$

Statistics for Singlet Full-CI Wave Function Expansions

$n=N$	N_{csf}	N_{row}	N_{ϕ}	N_{pair}^a	N_{value}^b	$t(H_{PQ})^c$	$t(E')^d$	$t(E';FD)^e$
2	3	5	2	10	26	0.00	0.00	0.00
4	20	14	13	43	278	0.00	0.00	0.00
6	175	30	39	120	1058	0.00	0.00	0.02
8	1,764	55	86	261	2682	0.00	0.01	0.10
10	19,404	91	160	486	5466	0.00	0.05	0.64
12	226,512	140	267	815	9726	0.00	0.16	3.20
14	2,760,615	204	413	1268	15778	0.01	0.44	12.39
16	34,763,300	285	604	1865	23938	0.04	1.24	36.24
18	449,141,836	385	846	2626	34522	0.07	3.48	118.44
20	5,924,217,936	506	1145	3571	47846	0.13	9.29	297.70
22	79,483,257,308	650	1507	4720	64226	0.21	25.67	632.94
24	1,081,724,803,600	819	1938	6093	83978	0.34	65.49	1.32E3
26	14,901,311,070,000	1015	2444	7710	107418	0.54	140.61	2.64E3
28	207,426,250,094,400	1240	3031	9591	134862	0.82	250.45	4.97E3
30	2,913,690,606,794,775	1496	3705	11756	166626	1.21	423.87	8.97E3
32	41,255,439,318,353,700	1785	4472	14225	203026	1.75	676.76	1.57E4
34	588,272,005,095,043,500	2109	5338	17018	244378	2.49	1.07E3	2.66E4
36	8,441,132,926,294,530,000	2470	6309	20115	290998	3.46	1.62E3	4.37E4
38	121,805,548,126,430,067,900	2870	7391	23656	343202	4.66	2.38E3	6.89E4
40	1,766,594,752,418,700,032,400	3311	8590	27541	401306	6.27	3.48E3	1.08E5
42	25,739,723,541,439,406,257,200	3795	9912	31830	465626	8.25	4.93E3	1.64E5
44	376,607,675,256,599,252,232,000	4324	11363	36543	536478	11.19	6.88E3	2.54E5
46	5,531,425,230,331,301,517,157,500	4900	12949	41700	614178	14.43	9.47E3	3.74E5

a) The number of node pairs that contribute to nonzero Shavitt loop values. This is also the number of vertices in the auxiliary pair graph data structure. b) The total number of nonzero segment values in the Shavitt graph. c) Times are in seconds on a 2.5GHz PowerMac G5 to construct a single

$\langle P|\hat{H}|Q\rangle$ matrix element. d) Times in seconds to construct the analytic gradient vector

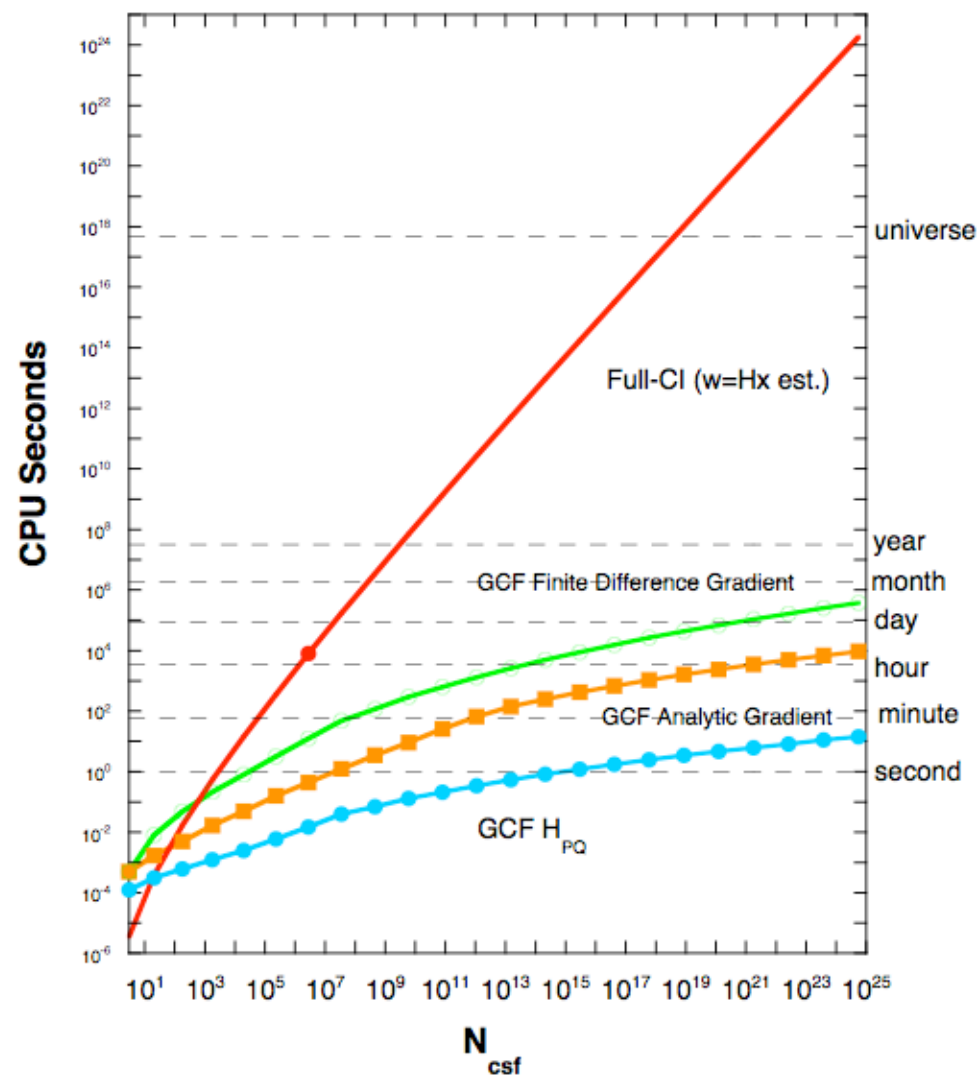
$E'(\phi_0) \equiv \partial E(\phi)/\partial \phi_{(mM)}|_{\phi_0}$ for $N_{\alpha}=1$ using the $\mathbf{G}^{[u]}$ and $\mathbf{S}^{[u]}$ arrays. e) Times in seconds to construct

the gradient with a finite-difference approximation, $t(E';FD)=2N_{\phi}t(H_{PQ})$.

GCF-CI...

$H_{PQ} = \langle P|H|Q \rangle$ and Gradient Construction Time
($n=N$, Singlet, Full-CI)

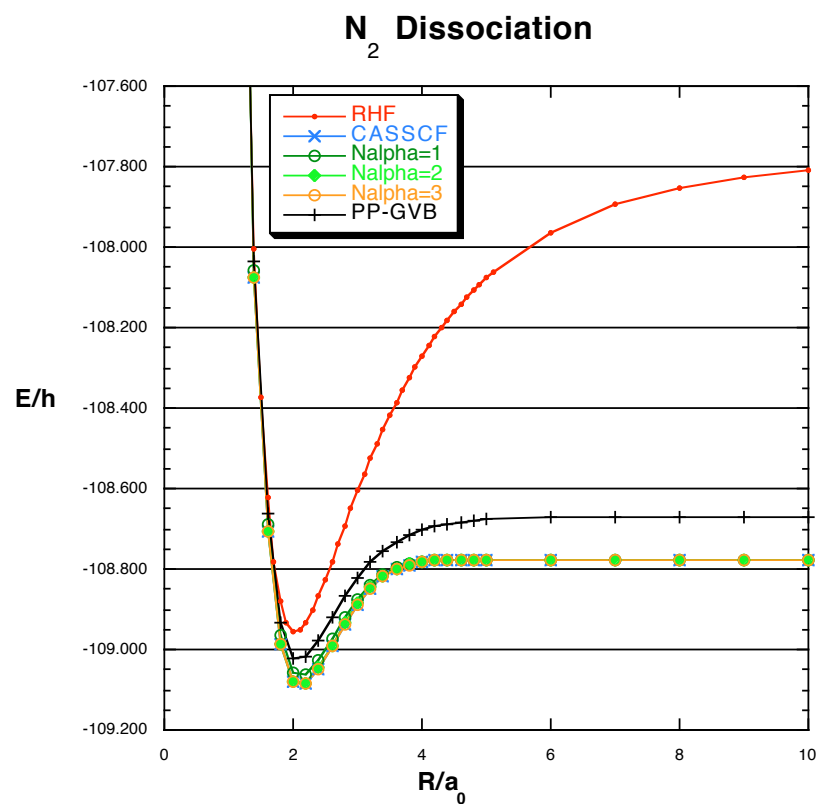
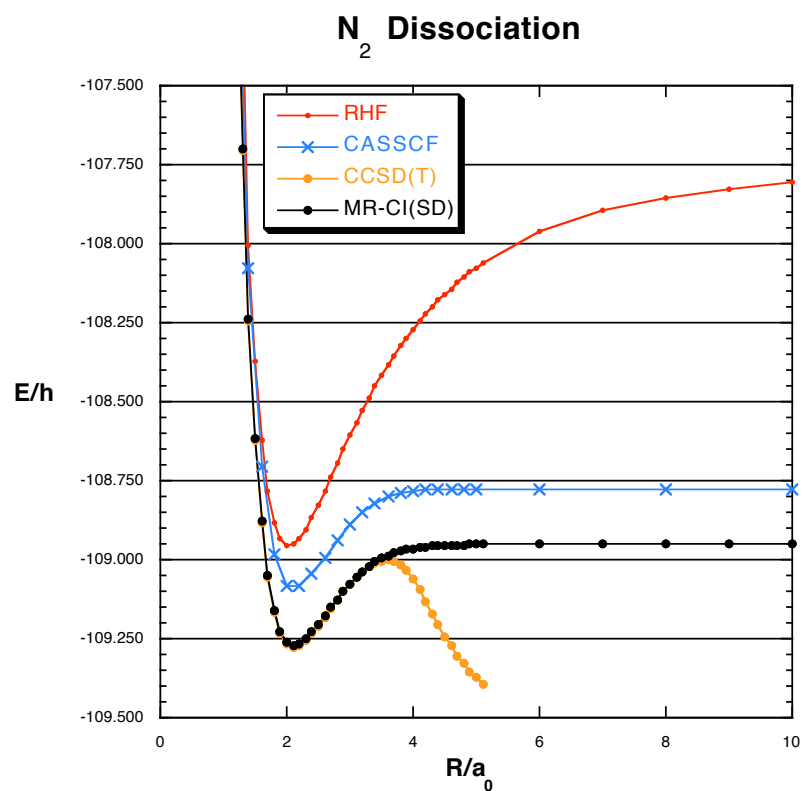
Quantity	Effort
S_{PQ}	$\mathcal{O}(N_{\text{row}}) = \mathcal{O}(\beta n)$
\mathbf{D}^{PQ}	$\mathcal{O}(\beta n^2)$
\mathbf{d}^{PQ}	$\mathcal{O}(\beta n^4)$
H_{PQ}	$\mathcal{O}(\beta n^4)$
H'_{PQ}	$\mathcal{O}(\beta n^5)$
Full-CI H_{PQ}	$\mathcal{O}(N^2 n^2 N_{\text{csf}}) \approx \mathcal{O}(n^2 4^n)$



GCF-Cl...

Valence Correlation:

N₂: 6⁶ Valence Full-CI



GCF-CI...

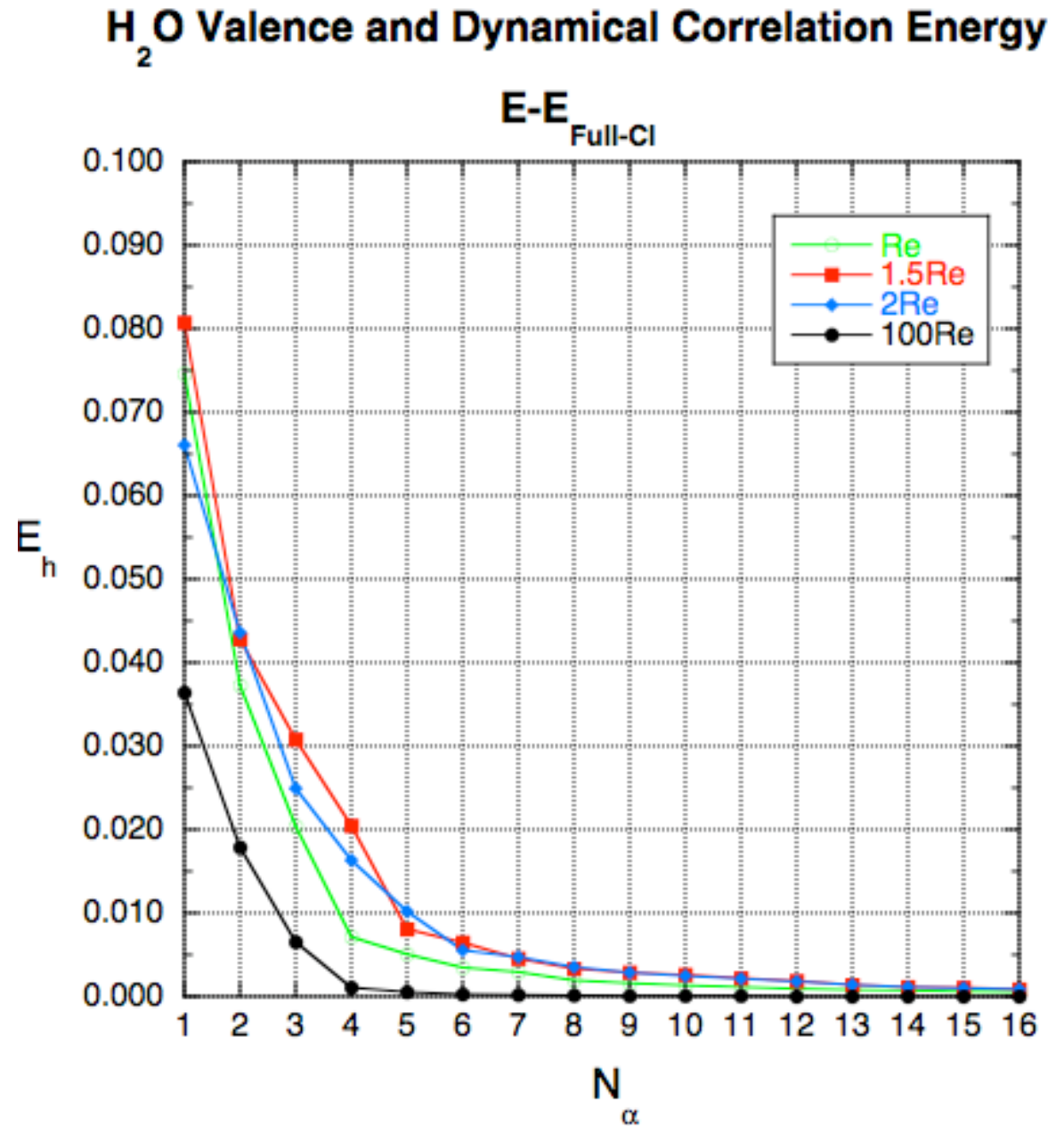
(14^{10}) Full-CI H_2O

$N_{\text{csf}}=256473$

$N_{\text{row}}=175$

$N_{\text{arc}}=514 \rightarrow 513$

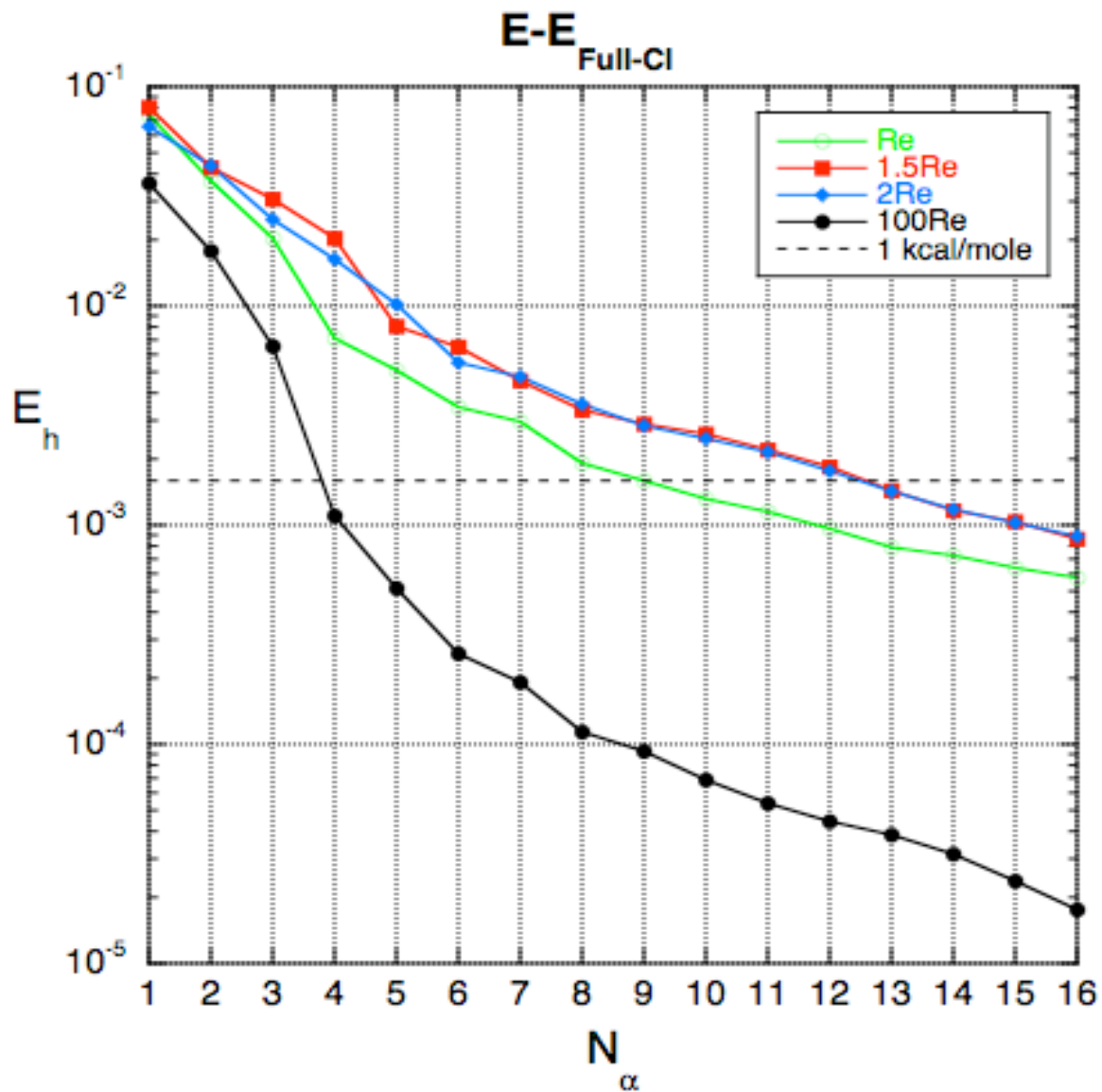
$N_{\phi}=340 \rightarrow 339$



GCF-CI...

(14^{10}) Full-CI H_2O

H_2O Valence and Dynamical Correlation Energy



GCF-Cl...

Wave Function Analysis Based on the Shavitt Graph – How to analyze a wave function expanded in 10^{100} CSFs?

- Node Density:
$$1 = \langle \psi | \psi \rangle = \sum_j^{(Level\ p)} D_j^\psi$$
$$= \sum_j^{(Level\ p)} \sum_{M,N} c_M c_N \gamma_j^{M,N} \bar{\gamma}_j^{M,N}$$
- Arc Density:
$$1 = \langle \psi | \psi \rangle = \sum_\mu^{(Level\ p)} D_\mu^\psi$$
$$= \sum_\mu^{(Level\ p)} \sum_{M,N} c_M c_N \alpha_\mu^M \alpha_\mu^N \gamma_{Bottom(\mu)}^{M,N} \bar{\gamma}_{Top(\mu)}^{M,N}$$

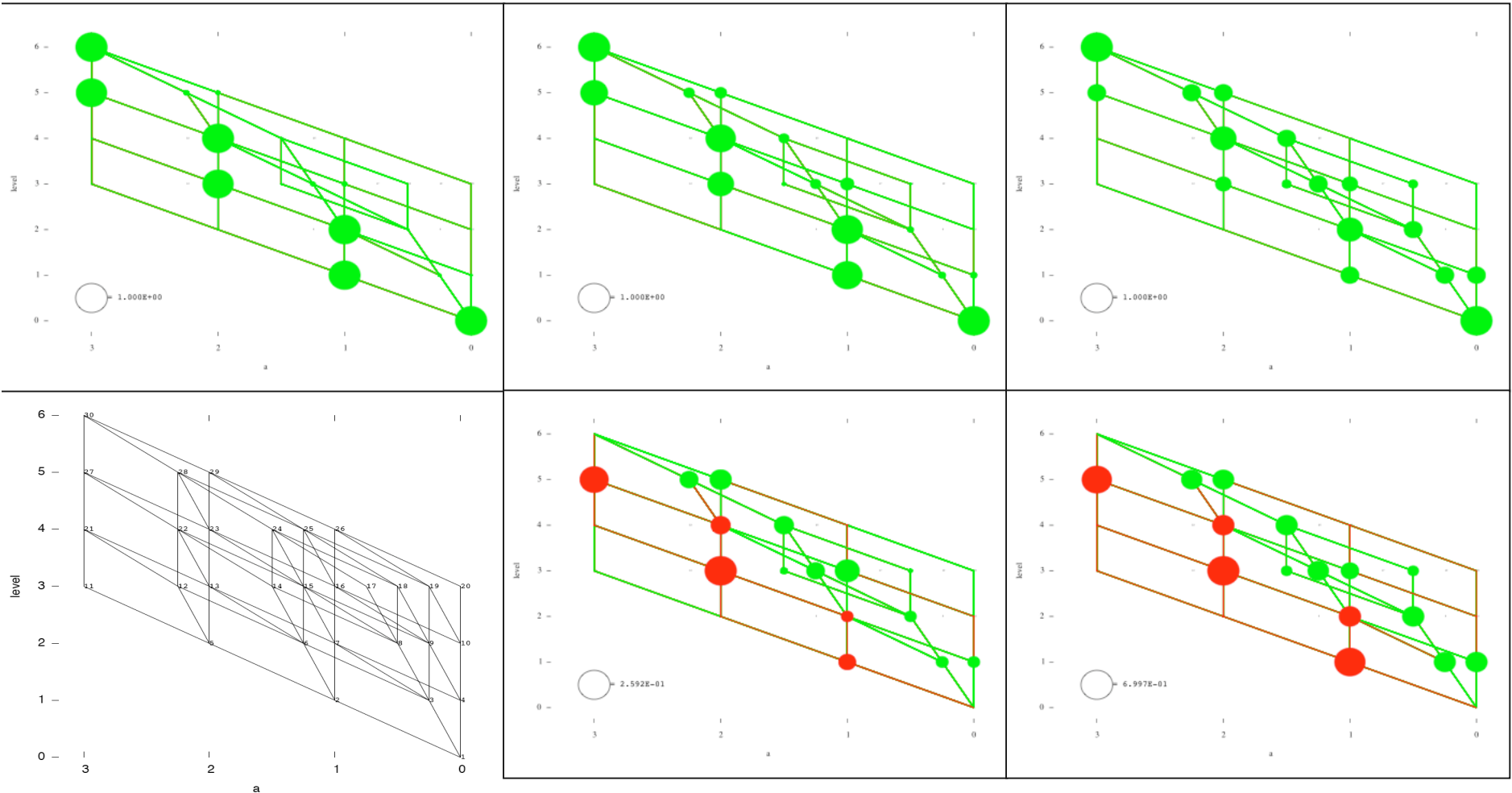
GCF-Cl...

N₂ Arc and Node Density

R=2.2

R=3.2

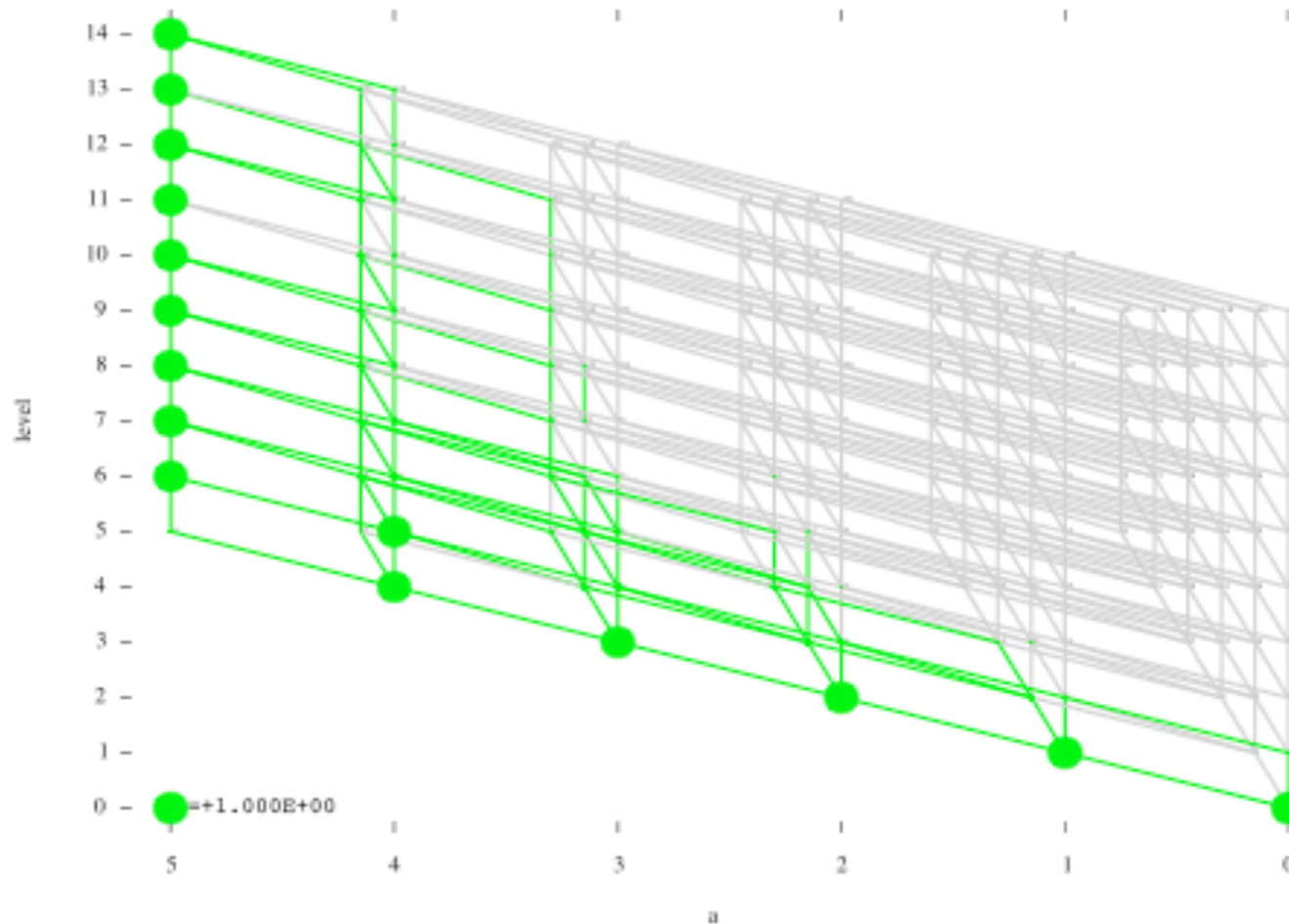
R=10.0



GCF-Cl...

H₂O Arc and Node Density

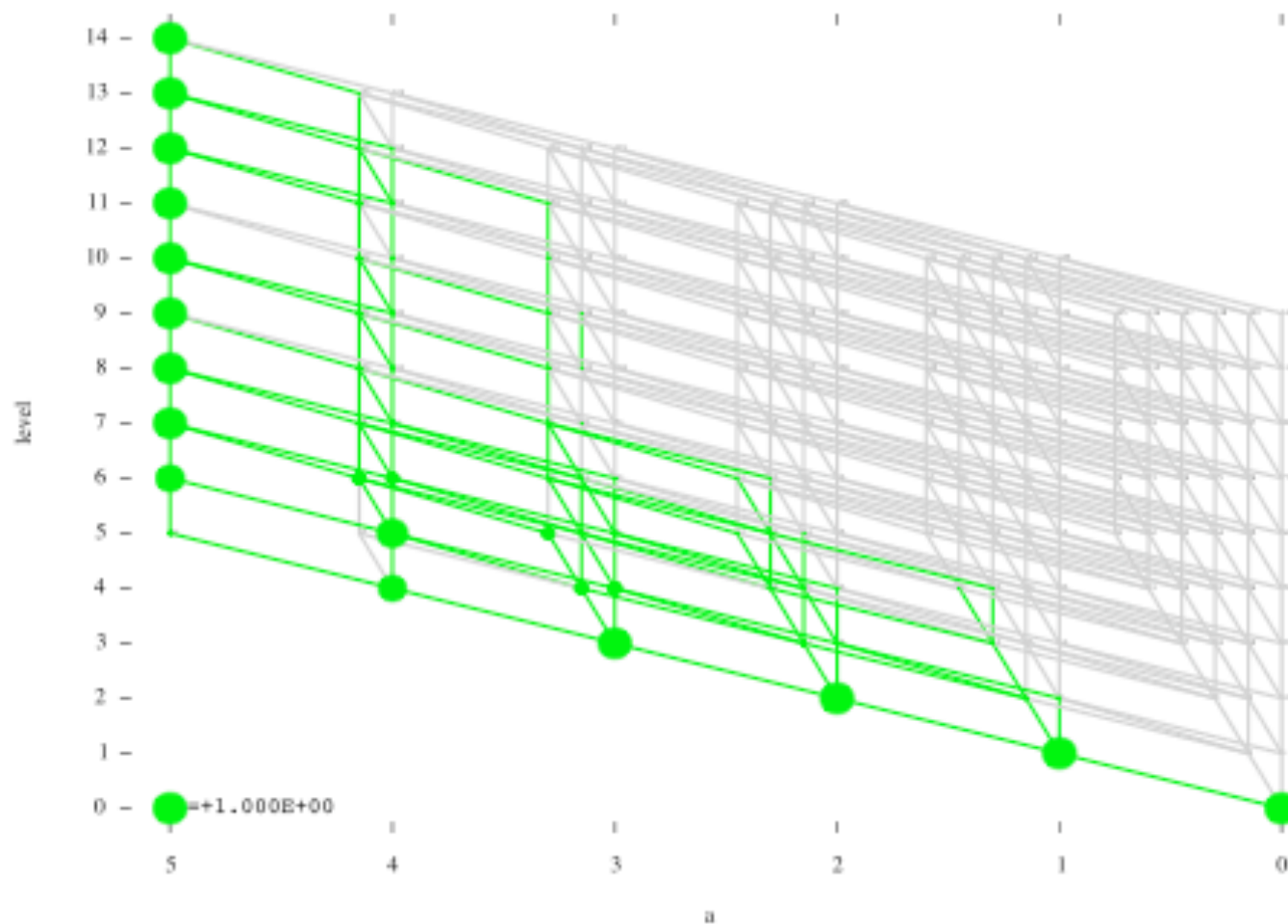
R_e Graph Density



GCF-Cl...

H₂O Arc and Node Density

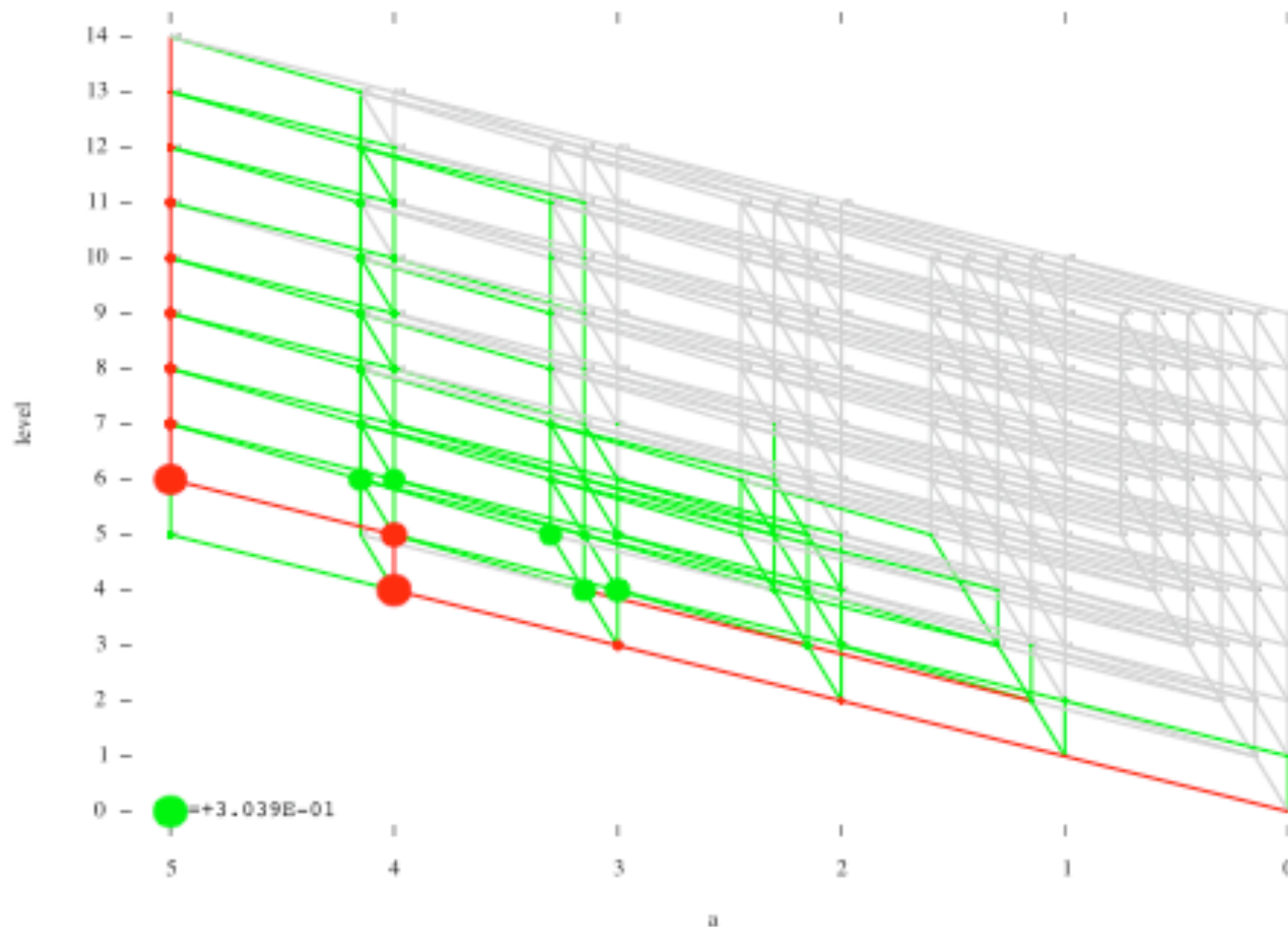
2R_e Graph Density



GCF-Cl...

H₂O Arc and Node Density

2R_e Density-R_e Density



Summary

- The GCF-CI approach has enormous potential because the S_{PQ} and H_{PQ} matrix element construction is very efficient.
- The storage requirements and the computational effort scale as N_{row} . No effort or storage scales as N_{csf} which can be, in principle, arbitrarily large.
- The method may be applied to both ground and excited states; the Ritz variational principle applies to all computed eigenvalues of $\mathbf{Hc}=\mathbf{ScE}$.
- Our implementation is based on GUGA (low- and high-spin states, no spin contamination, no spin instabilities).
- An initial effort has been made to allow wave function characterization based directly on the Shavitt Graph (arc density, node density, qualitative interpretation of the arc factors).

Summary...

- The wave function form is not based on expansion about a reference wave function; its accuracy is not limited or biased by a failure of the HF reference.
- There are no inherent excitation-level limitations (e.g. CI-SD, CC-SD(T), etc.). The wave function flexibility is determined by the underlying Shavitt graph.

This work was performed under the auspices of the Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, and SciDAC, U.S. Department of Energy, under contract number DE-AC02-06CH11357.